REMARKS

Applicant appreciates the time taken by the Examiner to review Applicant's present application. This application has been carefully reviewed in light of the Official Action mailed November 25, 2008. Applicant respectfully requests reconsideration and favorable action in this case.

Claims Status

Claims 1, 2, 13-15, 17, 21-22, 26, 30, 33, 40-41, 46-47, 50, 54, 56 and 59 were presented for examination. Claims 1, 2, 13-15, 17, 21-22, 26, 30, 33, 40-41, 46-47, 50, 54, 56 and 59 were rejected. Claims 1, 30, and 54 are amended herein. Support for the amendments may be found at least in paragraphs 63, 67-69, 110, 115 and 117 of the specification as filed. No new matter is added. Thus, claims 1-2, 13-15, 17, 21-22, 26, 30, 33, 40-41, 46-47, 50, 54, 56 and 59 remain pending.

Rejections under 35 U.S.C. § 103

Claims 1 and 54 were rejected as obvious over U.S. Patent No. 5,031,089 ("Liu") in view of U.S. Pre-Grant Publication No. 2005/0044206 ("Johannson"). The rejection is traversed. Claim 54 contains similar language as claim 1. Accordingly, the rejection will be traversed collectively as it pertains to claim 1.

Claim 1, as amended, recites:

A method for regulating resource usage by a plurality of distributed applications running on a plurality of interconnected machines, the method comprising:
providing a system comprising a plurality of machines connected to each other through a network, wherein the plurality of machines run one or more operating systems;
at each machine:
detecting applications running on the machine; and
detecting the operating system processes and network traffic associated with the application;
receiving an initial resource allocation policy based on the detected applications and operating system processes and network traffic associated with the applications;
gathering, by the plurality of interconnected machines, information on the state and health of applications and system resources;
distributing among the machines the initial resource allocation policy, wherein each machine receives a subset of the total information gathered and a portion of the initial resource policy, wherein the subset of the total information received by each machine corresponds to one or more applications running on that machine;
each machine detecting one or more requests for resources by one or more of

the plurality of distributed applications;
at each of the plurality of interconnected machines, allocating the resources to the one or more of the plurality of distributed applications based on the subset of the total information gathered;
determining the resource consumption by the each application in the plurality of distributed applications based on its actual usage;
periodically exchanging resource information amongst the plurality of interconnected machines, wherein the resource information includes requests for the resources, resource consumption, resource availability at each of the plurality of interconnected machines, the current demand levels and the application priorities; and
at each of the plurality of interconnected machines, adjusting the resources to each application of the plurality of distributed applications based upon the periodically exchanged resource information received by that machine.

Thus, embodiments of a method for regulating resource usage by a plurality of distributed applications running on a plurality of interconnected machines may include providing a system comprising a plurality of machines connected to each other through a network. At each machine, applications running on the machine can be detected along with the operating system processes and network traffic associated with the application. An initial resource allocation policy based on the detected applications and operating system processes and network traffic associated with the applications may be received by the system. The interconnected machines can gather information on the state and health of applications and system resources. The initial resource allocation policy can be distributed among the machines, wherein each machine receives a subset of the total information gathered and a portion of the initial resource policy. The subset of the total information received by each machine corresponds to one or more applications running on that machine. Each machine detects one or more requests for resources by one or more of the plurality of distributed applications and allocates its resources to the one or more of the plurality of distributed applications based on the subset of the total information gathered. The resource consumption by the each application in the plurality of distributed applications may be determined based on its actual usage, and the resource information (including requests for the resources, resource consumption, resource availability at each of the plurality of interconnected machines, the current demand levels and the application priorities) may be periodically exchanged amongst the plurality of interconnected machines. At each of the plurality of interconnected machines, the resources may be adjusted to each application of the plurality of distributed applications based upon the periodically exchanged resource information received by that machine.

Liu teaches a system for resource allocation in which resources are allocated to evenly distribute the system workload. The policy used by Liu appears to be such that when the workload for a node in the system falls below a predetermined level, the node checks with the other nodes to find the node with the highest workload level and then request some of the workload. In some cases, this is performed even when the node is idle. (See, Liu, Abstract.)

Johannson teaches a centralized system, in which a monitoring unit (31) monitors traffic to determine how to allocate resources. For example, if the monitoring unit detects a resource deficit for at least one category, it informs a reconfiguration unit (32) to calculate a new resource distribution policy on the basis of the monitored traffic. (See, Johannson, para. 26.)

In the rejection, the Examiner states that Liu discloses a method for regulating resource usage. Applicant respectfully submits that Liu is concerned instead with ensuring applications or other processes are being processed in a timely manner, independent of any resource allocation policy. For example, a node checks the workload of other nodes when its workload level falls below a predetermined level, and an idle node periodically checks other node workloads. Thus, it appears as though the driving factor in resource allocation is the workload level of the nodes. As Applicant mentioned, a drawback to this approach is that this can result in servers operating at 20% (or less) of capacity. (See, specification, para. 13.)

In contrast to the approach of Liu, embodiments disclosed by Applicant may be useful for ensuring servers are operating at higher capacity. Each machine receives a subset of the total information gathered and allocates resources according to the resource policy. Periodically, the interconnected machines exchange resource information, wherein the resource information includes requests for the resources, resource consumption, resource availability at each of the plurality of interconnected machines, the current demand levels and the application priorities (See, specification, para. 67). At each of the plurality of interconnected machines, the resources may be adjusted to each of the plurality of distributed applications based upon the periodically exchanged resource information. Thus, embodiments disclosed by Applicant may allocate resources based on a number of factors, not just whether a request has been received and there is an idle node or a node that has a greater capacity.

Applicant submits that the teachings of Johannson fail to remedy the shortcomings of Liu. Instead, Johannson teaches a system in which resources are allocated based on a maximum threshold. For example, in Johansson's description of dynamic resource allocation in packet-based communication networks, the monitoring unit keeps track of a set of counters for each of the various traffic categories and, if the counter for a category exceeds an allowed threshold value, then the monitoring unit informs the reconfiguration unit. (See, Johannson,

para. 26.) Thus, it appears as though Johannson waits until a threshold has been crossed before adjusting the allocation of bandwidth.

In contrast, embodiments disclosed by Applicant do not rely on monitoring to see if requests have exceeded a threshold, but instead periodically exchange information to see if the resources can be allocated (or reallocated) to improve data center efficiency while maintaining service level requirements of users.

Regarding claim 30, the Examiner states that Liu teaches a system for regulating utilization of computer resources and that Johannson teaches a manager module for providing a global distributed policy. In particular, the Examiner states that it would have been obvious to combine Liu and Johannson to set and establish the maximum values each resource can use on each of the interconnected machines. (See, Office Action mailed 06/01/2009, page 8, lines 2-3.) As mentioned above, embodiments disclosed by Applicant do not rely on monitoring to see if requests have exceeded a threshold, but instead periodically exchange information to see if the resources can be allocated (or reallocated) to improve data center efficiency while maintaining service level requirements of users.

For at least the foregoing reasons, Applicant respectfully submits that the allocation of resources based on minimum thresholds as taught by Liu combined with the centralized approach described by Johannson fails to teach a distributed system that periodically exchanges information and reallocates resources based on the periodically exchanged information disclosed by Applicant. Accordingly, withdrawal of this rejection is requested.

Regarding the rejections of claims 2, 13-15, 17, 21-22, 26, 33, 40-41, 46-47, 50, 56 and 59, reliance is placed on *In re Fine*. If an independent claim is nonobvious under 35 U.S.C. 103, then any claim depending therefrom is nonobvious. *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988). Claims 2, 13-15, 17, 21-22 and 26 depend from claim 1, claim 33, 40-41, 46-47 and 50 depend from claim 30, and claims 56 and 59 depend from claim 54, which Applicant believes to be non-obvious over Liu and Johannson. Accordingly, withdrawal of this rejection is requested.

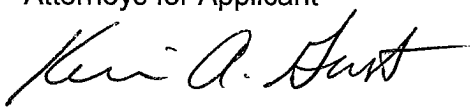
Conclusion

Applicant has now made an earnest attempt to place this case in condition for allowance. Other than as explicitly set forth above, this reply does not include an acquiescence to statements, assertions, assumptions, conclusions, or any combination thereof in the Office Action. For the foregoing reasons and for other reasons clearly apparent, Applicant respectfully requests full allowance of Claims 1-2, 13-15, 17, 21-22, 26, 30, 33, 40-41, 46-47, 50, 54, 56 and 59. The Examiner is invited to telephone the undersigned at the number listed below for prompt action in the event any issues remain.

The Director of the U.S. Patent and Trademark Office is hereby authorized to charge any fees or credit any overpayments to Deposit Account No. 50-3183 of Sprinkle IP Law Group.

Respectfully submitted,

Sprinkle IP Law Group
Attorneys for Applicant



Kevin A. Gust
Patent Agent
Reg. No. 51,032

Date: September 1, 2009

1301 W. 25th Street, Suite 408
Austin, TX 78705
Tel. (512) 637-9220
Fax. (512) 371-9088